Screening of a Novel Fragment Library with Functional Complexity against *Mycobacterium tuberculosis* InhA


Our findings reported herein provide support for the benefits of including functional group complexity (FGC) within fragments when screening against protein targets such as *Mycobacterium tuberculosis* InhA. We show that InhA fragment actives with FGC maintained their binding pose during elaboration. Furthermore, weak fragment hits with functional group handles also allowed for facile fragment elaboration to afford novel and potent InhA inhibitors with good ligand efficiency metrics for optimization.

We recently reported on the design and synthesis of a novel fragment library of diverse fragments which include functional group complexity (FGC). Functional groups, or “chemical handles”, were incorporated onto diverse scaffolds to allow for additional interactions with target proteins with the aim of maintaining binding poses during optimization, as well as aid fragment elaboration. However care has to be taken because increasing complexity in a fragment decreases the probability of it achieving optimal ligand-protein interactions. Conversely, too little complexity can lead to interesting interactions being missed. In this respect, fragment deconstruction studies on β-lactamase inhibitors, suggest that small fragments with “minimal complexity” derived from the deconstruction of potent inhibitors do not always retain the binding poses of the parent molecules. Whereas, fragments with built in FGC recapitulate the larger potent inhibitor binding mode. Therefore, a careful balance is required to identify a sweet spot of complexity where detectable, single-mode binding of a ligand to a target is most probable. There is further support for this concept based on other fragment deconstruction case studies.

Despite the molecular complexity model and its putative applications in fragment screening were introduced and refined by Hann and others over nearly two decades ago, to date the authors are not aware of any reported fragment screening against a novel diverse fragment library designed to include FGC. Our recently reported FGC library is based on synthetic chemistry toward selected functional groups. However, recent reports of an algorithm to identify all functional groups in organic molecules, allows for the analysis of FGC in large chemical databases of commercial fragments.

Herein we describe the screening of such a library, in comparison with other fragment sets, against the highly validated *Mycobacterium tuberculosis* (Mtb) target InhA. Mtb is the causative agent of tuberculosis (TB), which is currently the leading infectious disease killer worldwide. Isoniazid (INH, Figure 1), a successful frontline TB drug for more than 50 years, targets the NADH-dependent 2-trans enoyl–acyl carrier protein

![Figure 1. INH and selected advanced direct InhA inhibitors 1–3.](https://doi.org/10.1002/cmdc.201700774)
(ACP) reductase InhA. This is a key enzyme in the Mtb cell wall synthesis pathway\(^{[10]}\) and does not have a human orthologue. The development of Mtb resistant strains to standard anti-TB drugs,\(^{[11]}\) including INH, necessitates the need for novel Mtb targeted therapies. Resistance to INH develops mainly via mutations in the Mtb KatG enzyme, which converts INH into an acyl radical, which covalently binds to NADH and the resulting adduct inhibits InhA.\(^{[12]}\) Direct InhA inhibitors are envisaged to bypass this resistance mechanism and maintain clinical efficacy. Accordingly, there has been widespread research in this field\(^{[13]}\) and whilst limited set of potent direct InhA inhibitors with activity against INH-resistant strains have been identified (1-3\(^{[14]}\) in Figure 1), none have been progressed into clinical development. Hence, there remains a need to identify novel direct InhA inhibitor scaffolds.

InhA inhibitors are known to modulate the tertiary structure of the InhA protein binding pocket, in particular the substrate binding loop (SBL).\(^{[15]}\) In this respect, a fragment based (FB) approach\(^{[16]}\) was considered appealing in order to assess the InhA protein conformations for fragment actives and the structural requirements for their optimization into potent InhA inhibitors.

For the above reasons, we screened the recently reported FGC fragment library (FGC-FRAG),\(^{[1]}\) as well as an informed InhA commercial fragment set (InhA-INF-FRAG), which was compiled based on the known direct InhA inhibitors in the public domain (see Supporting Information). The above libraries were screened alongside an historical commercial fragment library (HIST-FRAG), a reported 3D fragment library (3D-FRAG)\(^{[17]}\) and fragments derived from inventory (INV-FRAG) and project (PROJ-FRAG) sources. The overall library constituted 1360 fragments (Figure 2A), which were screened against the NADH bound form of the InhA, using saturation transfer difference (STD) \(^1\)H NMR (complete results in Supporting Information).

STD-NMR typically identifies ligands that bind weakly to moderately to protein targets.\(^{[18]}\) The criteria for a binding event used here was a positive STD signal intensity which was decreased by at least 50 % on the addition of the known inhibitor 1 (R = Me).\(^{[14]}\) This resulted in 149 hits (11 % hit rate). A breakdown of these hits based on their source is given in Figure 2B. Due to its binding affinity being in the suitable range (K_d \approx 5 \mu M),\(^{[19]}\) NADH binding was also observed in the STD-NMR spectra. It was noted that the stronger binders 1 (R = Me) and 3 (R = CH_3Pr)\(^{[14b]}\) decreased the STD-NMR intensities for the NADH co-factor peaks. Therefore, greater importance was given to those fragments which also caused a decrease in the NADH STD peak intensities, as this was considered as evidence of stronger binding. This further selection step decreased the number of hits to 32 (4-35 in Figure 3; 2.4 % hit rate). The pie chart for the source of these 32 hits is given in Figure 2C. This process increased the fraction of hits from the FGC-FRAG set (29 % to 41 %). These data are interesting considering the FGC-FRAG set only constituted 24 % of the whole screening library. The initial hit rate for the InhA-INF-FRAG set was low, although the size of the library was small. This may be the result of a lack of InhA fragment inhibitors that can be purchased from vendors, as observed for deconstruction of kinase inhibitors from the public domain.\(^{[20]}\) The two hits derived from this library did, however, survive the second selection step and could also be classified as FGC fragments. A high proportion of project, historical derived and 3D fragment actives were also noticeably enriched with FGC.

The 149 NMR hits were also screened in a high concentration (500 \mu M) biochemical assay. Only fragments 4 (13 %), 9 (37 %), 22 (11 %), and 34 (10 %) showed InhA inhibitory potencies < 10 %, and were further tested in dose–response studies up to 1 mM (Table 1). Notably, these four fragment hits were all from the 32 compounds expected to be more potent based on NADH STD signal suppression. Based on these results as well as chemical diversity, 15 compounds were prioritized for crystallography studies. Crystals suitable for structure determination were obtained for fragments 4, 9, 12, 22, 24 and 34 bound to InhA.

Surface plasmon resonance (SPR) equilibrium dissociation binding constants (K_d) were also determined, and found to correlate with InhA biochemical pIC_{50} values (Table 1).
The InhA–NADH co-crystal structures for FGC fragments 22 (pIC$_{50}$ = 3.1) and 24 (pIC$_{50}$ < 3) were found to be in accordance with the FGC-FRAG set design principle, in that additional interactions were observed for their FG with the InhA protein (Figures S1 and 4A). The co-crystal structure of 24 contains a unique tetramer with four slightly different binding sites. The following description is for one of these, details of the others are shown in Figure S2. The amide carbonyl FG of 24 forms an H-bond interaction with the InhA Tyr158 residue and the NADH ribosyl-20-OH group, while the NH H-bonds with Met199. Furthermore, the pyrimidine ring of 24 picks up an additional H-bond interaction with the backbone NH of Met98, a feature also observed for the thiadiazole core of a close analogue (PDB ID: 4BQP) of advanced lead 2, which, however, does not interact with Tyr158. Whereas lead 3a (PDB ID: 4R9S) has a similar H-bond interaction with Tyr158 and the NADH ribosyl-20-OH group. FGC fragment 24 is able to identify distinct interactions associated with both leads 2 and 3a (Figure 4A).

Fragment 24 was found to overlay well with HIST-FRAG urea 12 (pIC$_{50}$ < 3.0), with the embedded urea FG having similar interactions with Tyr158, Met199 and the NADH ribosyl hydroxy group (Figure 4A). Furthermore, the pyridyl group of 12 occupies the same space as the lipophilic di-methyl cyclohexane group in lead 3a (Figure S3), however, its nitrogen is predicted to be protonated and binds to the carboxylic moiety of Glu219. Overlay of the urea 12 with the amide 24 led to synthesis of the merged amide 36 and urea 37 (Figure 5 and Scheme S1). Amide 36 was inactive in the biochemical screen (pIC$_{50}$ < 3), but urea 37 had improved InhA biochemical poten-

Figure 3. 32 STD NMR hits 4–35. FGC are in blue; known InhA cores are in red. Crystal structures were obtained for fragments in bold.
Preparation of individual fragment components was conserved (Figure 4 B). ChemMedChem cy (pIC 50
3a both the advanced leads of 40 and ring of NADH, and the 2-N
mediated bridge with NADH. However, the pyrazole rings (pIC50
action with the 20-OH group of NADH.
structures obtained for pyrazole fragment hits 4
led to the identification of pyridinone (pIC50
ferred onto the pyrimidine series through synthesis of urea
ment of the phenyl group of 41 overall pose of the individual FGC fragment components of
with improved enzymatic activity (pIC 50
fragment afforded compound 45 (PDB ID: 5JFO, Figure 6 B). The 3N of the thiazazole ring and NH side group of
and 45 H-bonds with Met98 backbone.
Because good structural overlay (Figure 56) was observed for 9 and the published InhA lead 44 (pIC50=6.9, PDB ID: 4QXM), compound 46 (Figure 7 and Scheme S4) was synthesized and showed good InhA activity (pIC50=6.3). Introducing the amide FG results in movement of the phenyl group of 9 to allow an additional H-bond interaction with Met98 backbone CO (Figure 6C). Similarly, as a result of the deconstructed fragment 34 retaining the binding pose of 45 (pIC50=8.1) as well as occupying the same pyrazole binding pocket as fragment 9, compound 47 (Figure 7 and Scheme S5) was synthesized. 47 showed good InhA activity (pIC50=7.3) and maintained the pose relative to 45 (Figure S7).
In conclusion, herein we report on the identification of novel InhA fragment hits using STD-NMR screening, as well as orthogonal InhA biochemical and SPR assays. High hit rates were obtained from screening a recently reported novel fragment set with built FGC versus other fragment sets. Notably, starting from weakly active FGC fragments facilitated rapid fragment based lead generation (FBLG) due to 1) easy chemical tractability and derivatization, 2) retention of the functional group binding pose during fragment evolution through additional interactions with the target, which was confirmed by X-ray studies. Conversely, elaboration of weakly bound InhA frag-

### Table 1. Biochemical, SPR and LE metrics for fragment hits and optimized InhA inhibitors.

| Compound | InhA pIC50 [\textsuperscript{exo}] (\% 500 \mu M) | PDB ID | SPR pK\textsubscript{A} [\textsuperscript{exo}] | InhA biochemical LE/LLE/LELP [\textsuperscript{a}]
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (R = Me)</td>
<td>7.9 ± NA</td>
<td>–</td>
<td>ND</td>
<td>0.30/5.0/5.4</td>
</tr>
<tr>
<td>3 (R = CH3P)</td>
<td>5.0 ± 0.7\textsuperscript{[b]}</td>
<td>–</td>
<td>4.4 ± 0.0\textsuperscript{[c]}</td>
<td>0.35/2.3/7.8</td>
</tr>
<tr>
<td>4</td>
<td>&lt; 3 (13%)</td>
<td>5OIC</td>
<td>NA\textsuperscript{[d]}</td>
<td>–</td>
</tr>
<tr>
<td>9</td>
<td>3.4 ± 0.23 (37%)</td>
<td>5OF</td>
<td>3.3 ± 0.08\textsuperscript{[e]}</td>
<td>0.34/3.1/5.1</td>
</tr>
<tr>
<td>12</td>
<td>&lt; 3 (2%)</td>
<td>5OL</td>
<td>NA\textsuperscript{[f]}</td>
<td>–</td>
</tr>
<tr>
<td>22</td>
<td>3.1 ± 0.05 (11%)</td>
<td>5OIM</td>
<td>NA\textsuperscript{[g]}</td>
<td>0.28/0.5/9.3</td>
</tr>
<tr>
<td>24</td>
<td>&lt; 3 (0%)</td>
<td>5OIN</td>
<td>NA\textsuperscript{[h]}</td>
<td>–</td>
</tr>
<tr>
<td>34</td>
<td>&lt; 3 (10%)</td>
<td>5OIO</td>
<td>ND</td>
<td>–</td>
</tr>
<tr>
<td>37</td>
<td>&lt; 3</td>
<td>–</td>
<td>ND</td>
<td>–</td>
</tr>
<tr>
<td>38</td>
<td>4.1 ± 0.06</td>
<td>5OIP</td>
<td>NA\textsuperscript{[i]}</td>
<td>0.24/3.4/6.2</td>
</tr>
<tr>
<td>39</td>
<td>4.0 ± 0.01</td>
<td>–</td>
<td>3.2 ± 0.10\textsuperscript{[j]}</td>
<td>0.30/1.7/7.5</td>
</tr>
<tr>
<td>42</td>
<td>5.0 ± 0.14</td>
<td>5OIR</td>
<td>4.7 ± 0.03\textsuperscript{[k]}</td>
<td>0.30/4.2/6.5</td>
</tr>
<tr>
<td>49</td>
<td>4.8 ± 0.20</td>
<td>5OIL</td>
<td>4.6 ± 0.02\textsuperscript{[l]}</td>
<td>0.45/2.6/5.2</td>
</tr>
<tr>
<td>50</td>
<td>6.0 ± 0.06</td>
<td>5OIO</td>
<td>5.2 ± 0.02\textsuperscript{[m]}</td>
<td>0.39/4.4/6.9</td>
</tr>
<tr>
<td>51</td>
<td>6.5 ± 0.03</td>
<td>–</td>
<td>NA\textsuperscript{[n]}</td>
<td>0.33/4.6/13</td>
</tr>
<tr>
<td>52</td>
<td>5.2 ± 0.06</td>
<td>–</td>
<td>3.7 ± 0.10\textsuperscript{[o]}</td>
<td>0.30/3.4/10</td>
</tr>
<tr>
<td>57</td>
<td>6.9 ± 0.07</td>
<td>5OIP</td>
<td>NA\textsuperscript{[p]}</td>
<td>0.36/3.1/10</td>
</tr>
<tr>
<td>45</td>
<td>8.1 ± 0.27</td>
<td>4QXM</td>
<td>NA\textsuperscript{[q]}</td>
<td>0.40/4.7/9.4</td>
</tr>
<tr>
<td>46</td>
<td>6.3 ± 0.25</td>
<td>5OIS</td>
<td>6.3 ± 0.10\textsuperscript{[r]}</td>
<td>0.32/9.9/8.8</td>
</tr>
<tr>
<td>47</td>
<td>7.3 ± 0.20</td>
<td>5OIT</td>
<td>7.1 ± 0.06\textsuperscript{[s]}</td>
<td>0.36/5.1/8.2</td>
</tr>
</tbody>
</table>

[a] Compounds were tested up to 1 nM. (b) Compounds were tested up to 100 \mu M. (c) LE metrics were calculated using Stradrop in silico prediction software. (d) Data previously reported.\textsuperscript{[14c]} (e) Data are the mean ± SD of one independent experiment performed in duplicate. (f) Data are the mean ± SD of two independent experiments, each performed in duplicate. (g) No \( K_i \) values are quoted, but some interaction was observed. ND = not determined, NA = not available.
Figure 4. Crystal structures showing novel InhA–NADH–ligand complexes: A) overlay for fragments 12 (yellow) and 24 (blue), B) merged urea lead 37, and C) overlay for fragment 40 (blue) and fused lead 41 (yellow).

Figure 5. FBLG strategies: merging and growing.

Figure 6. Crystal structures showing novel InhA–NADH–ligand complexes: A) overlay for fragments 4 (blue) and 9 (yellow), B) overlay of 34 (yellow) with the published advanced lead 45 (blue; PDB ID: 5JFO), and C) overlay for fragment 9 (yellow) and lead 46 (blue).

Figure 7. Design strategy for pyrazoles 46 and 47.
ments with minimal FGC, relies on FGC implementation, which is likely to alter binding conformation.[23] This is a common issue associated with fragment optimization, resulting in structure–activity relationship disconnections which are often difficult to interpret. Our findings are also in agreement with the molecular complexity theory by Han et al.,[13,24] for which moderately complex ligands, like the identified FGC fragment hits, have a higher probability of a “useful event”, that is the detection of a unique binding pose. These results reported here provide support for the rational design, synthesis and screening of novel diverse fragments with built in functional groups. The described InhA FB-leads showed good InhA enzymatic activity as well as ligand efficiency (LE) metrics.[25] Additional optimization efforts have resulted in further improved InhA biochemical as well as Mtbo whole cell potency, which will be reported elsewhere.

Acknowledgements

We thank Laura Cleghorn and Penelope Turner for the synthesis of compounds 1 and 3. We thank David Robinson for compound handling support. We thank Ian Gilbert and Rob Young for helpful discussions and feedback. The research leading to these results was co-funded by the Tres Cantos Open Lab Foundation (project code TC162) and the European Union’s Seventh Framework Programme (grant agreement no. 291799). Fragment (project code TC162) and the European Union’s Seventh Framework Programme (grant agreement no. 291799). Fragment screening was funded by a MRC-CiNC grant (no. MC_PC_13061) to the University of Dundee.

Conflict of interest

The authors declare no conflict of interest.

Keywords: fragment based drug discovery · functional group complexity · InhA · tuberculosis


Manuscript received: December 11, 2017
Accepted manuscript online: February 5, 2018
Version of record online: 11 11 0000

These are not the final page numbers!
Functional group complexity fragment hits facilitated rapid fragment-based lead generation against *Mycobacterium tuberculosis* InhA, to afford novel and potent InhA inhibitors with good ligand efficiency metrics for optimization.